

Customer Number  
00909

Application Serial No.: 10/042,192  
Attorney Docket No.: 042846-0312966  
Pre-Appeal Brief Request for Review



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT(S) : Thangaraj Veerappan

CONFIRMATION 6876  
No.:

SERIAL NUMBER 10/042,192

EXAMINER : Schlaifer,  
Jonathon D.

FILING DATE : January 11, 2002

ART UNIT : 2178

FOR : SYSTEM AND METHOD FOR DETERMINING A DOCUMENT LANGUAGE AND  
REFINING THE CHARACTER SET ENCODING BASED ON THE DOCUMENT  
LANGUAGE

Mail Stop AF  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

***Pre-Appeal Brief Request for Review***

Sir:

In response to the Final Office Action mailed June 9, 2005, Applicant requests review of the final rejection in the above-identified application. This request is being concurrently filed with a Notice of Appeal. The review is requested for the reasons provided in the **Remarks** beginning on page 2 of this paper. A total of 5 pages are provided.

It is not believed that extensions of time or fees for net addition of claims are required beyond those that may otherwise be provided for in documents accompanying this paper. However, if additional extensions of time are necessary to prevent abandonment of this application, then such extensions of time are hereby petitioned for under 37 C.F.R. § 1.136(a), and any fees required therefore (including fees for net addition of claims) are hereby authorized to be charged to our Deposit Account No. 033975 (**Ref. No. 042846-0312966**).

## **Remarks**

Claims 1-40 are pending in this application. No claims are amended, added, or canceled by this paper. In view of the following remarks, allowance of all the claims pending in the application is requested.

### **Rejection Under 35 U.S.C. § 103**

Claims 1-4, 6, 11-14, 16-24, 26-34, and 36-40 stand rejected under 35 U.S.C. 103(a) as allegedly being unpatentable over U.S. Patent No. 6,252,671 to Peng *et al.* ("Peng") in view of U.S. Patent No. 6,104,381 to Watanabe *et al.* ("Watanabe"). Applicant traverses this rejection on the basis that the references cited do not teach or suggest all of the features of the claimed invention.

Claim 1 recites, among other things, evaluating at least a portion of the character string by comparing each of the characters in the portion of the character string to a plurality of predetermined candidate character sets to determine one or more matches between the plurality of pre-determined candidate character sets and the character string. Independent claims 11, 21, and 31 include similar recitations, among other things.

The invention is directed to receiving a character string in an unknown language and determining a language used to create the character string by performing a character-by-character analysis of the character string. In an exemplary embodiment, a textual message may be tested against a plurality of character sets by creating a bit mask from each of the character sets and parsing the textual message using the bit masks to return logical values that may flag the character sets that survive the character-by-character scan (see the Specification at page 9, line 23 – page 10, line 6). The logical values may be used to determine matches between the characters in the textual message and the character sets (see the Specification at page 10, lines 8 – 14).

Peng, on the other hand, appears to disclose an apparatus that downloads a font available on a computer to a printer that supports a page description language such as PostScript (see Peng at col. 2, lines 15 – 19). Thus, in Peng, information related to a

font, such as language information, would be included in font data transmitted between the computer and the printer in the page description language. Peng teaches that the apparatus enables both horizontal and vertical fonts to be downloaded, and facilitates the downloading of fonts that include more than 16K glyphs while reducing unnecessary processing and data storage requirements on the printer (see Peng at col. 2, line 21 – col. 3, line 14).

The Examiner contends that Peng “identifies character set encoding in the font documents in col. 5, lines 40-65,” and that this character set encoding is analogous to evaluating at least a portion of the character string by comparing each of the characters in the portion of the character string to a plurality of predetermined candidate character sets to determine one or more matches between the plurality of pre-determined candidate character sets and the character string. See the Final Office Action at paragraph number 6. While the cited passage appears to teach character set encoding to support various fonts, the assertion that this encoding is analogous to the claimed feature is patently incorrect. More particularly, the cited portion of Peng is directed to character set encoding appears to teach determining if a font includes primarily English or non-English symbols. If the font includes primarily non-English symbols, the font may either be encoded by either assigning each of the symbol characters within the font to a number (e.g., between 0 and 255), or by implementing a predetermined, default encoding scheme. Once the font is encoded, a symbol within the font may be specified in communications transmitted in a page description language, between, for example, a printer and a computer, by merely providing the number that corresponds to the symbol. However, the cited portion of Peng does not teach or suggest evaluating at least a portion of the character string by comparing each of the characters in the portion of the character string to a plurality of predetermined candidate character sets to determine one or more matches between the plurality of pre-determined candidate character sets and the character string.

The Examiner acknowledges that Peng is deficient *at least* for failing “to disclose that the document includes a character string, wherein the characters in the character string are represented in at least one of a plurality of character sets correspond to an

undermined language.” See *id.* The Examiner relies on Watanabe for this feature. Watanabe apparently describes an apparatus for inputting characters that may be printed onto a tape. See Watanabe at col. 1, lines 6-11. Even if Watanabe did disclose character strings, and assuming the existence of proper motivation for combining Watanabe with Peng, the proposed combination would still not resolve the deficiency of Peng addressed above because, as is the case with Peng, Watanabe does not teach or suggest evaluating at least a portion of the character string by comparing each of the characters in the portion of the character string to a plurality of predetermined candidate character sets to determine one or more matches between the plurality of pre-determined candidate character sets and the character string.

For *at least* these reasons, the rejection of claim 1 must be withdrawn. Independent claims 11, 21, and 31 are allowable over Peng at least for similar reasons. Claims 2-4, 6, 12-14, 16-20, 22-24, 26-30, 32-34, and 36-40 depend from and add additional features to corresponding ones of independent claims 1, 11, 21, and 31. Therefore, claims 2-4, 6, 12-14, 16-20, 22-24, 26-30, 32-34, and 36-40 are allowable, by virtue of their dependency, as well as for the features that they add to the independent claims.

Claims 5, 15, 25, and 35 stand rejected under 35 U.S.C. 103(a) as allegedly being unpatentable over Peng in view of Watanabe, and in further view of Schulze (U.S. Patent No. 6,167,369). Applicants traverse this rejection on the following basis.

The Examiner admits that the proposed combination of Peng and Watanabe fails to disclose that the at least one group of characters is an n-gram, but alleges that Schulze teaches that n-grams may be used to facilitate probabilistic analysis of whether a language is predominant (see the Office Action at paragraph 43). Schulze is apparently drawn to automatically identifying a predominant language of a document based on both probability data and word data. However, Schulze does not teach or suggest the deficiencies of Peng in view of Watanabe discussed above. Therefore, even if Examiner’s allegation with respect to Schulze is accepted, claims 5, 15, 25, and 35 are distinguishable over the cited references at least because the proposed

Customer Number  
00909

Application Serial No.: 10/042,192  
Attorney Docket No.: 042846-0312966  
Pre-Appeal Brief Request for Review

combination of Peng, Watanabe, and Schulze fail to provide the features of independent claims 1, 11, 21, and 31.

If, for any reason, personal communication will expedite prosecution of this application, the Examiner is invited to telephone the undersigned at the number provided.

Dated: August 25, 2005

Respectfully submitted,

A handwritten signature in black ink, appearing to read "Sean L. Ingram", is written over a horizontal line.

Sean L. Ingram  
Registration No.: 48,283  
PILLSBURY WINTHROP SHAW PITTMAN LLP  
1650 Tysons Blvd.  
McLean, Virginia 22102  
703-905-2000



## Addendum

### Invention Title

SYSTEM AND METHOD FOR DETERMINING A DOCUMENT LANGUAGE AND  
REFINING THE CHARACTER SET ENCODING BASED ON THE DOCUMENT  
LANGUAGE